**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

Graphical user interface, text

Description automatically generated

Μ = 0.3327 , σ = 0.1694 ,σ2=0.0287

For outliers we describe the data and the plot its boxplot

Graphical user interface

Description automatically generated with medium confidence

From the boxplot it is clear that the data is positively skewed and has outlier on the upper side which is at 91.36%



Answer the following three questions based on the box-plot above.

1. What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

Ans:- The Inter-Quartile Range (IQR) of the dataset is Upper Quartile – Lower Quartile

12 -5 = 7 approximately.

1. What can we say about the skewness of this dataset?

Ans:- The data is positively skewed as the mean is nearer to the minimum and also there is a outlier on the maximum side of the data.

1. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

Ans:- The data will have no outlier and also there will be slight changes in the median and the skewness of the data .



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans:- The mode of the data will lie approximately in between 4 to 8

1. Comment on the skewness of the dataset.

Ans :- The Data is Positively Skewed as the distribution curve drawn from the histogram will have large tail going to higher end .

1. Suppose that the above histogram and the box-plot in question 2 are plotted for the same dataset. Explain how these graphs complement each other in providing information about any dataset.

Ans :- From the boxplot we could easily find the outlier and with the help of Histogram we can see weather the data is normally distributed or not by drawing the Distribution Curve .

1. AT&T was running commercials in 1990 aimed at luring back customers who had switched to one of the other long-distance phone service providers. One such commercial shows a businessman trying to reach Phoenix and mistakenly getting Fiji, where a half-naked native on a beach responds incomprehensibly in Polynesian. When asked about this advertisement, AT&T admitted that the portrayed incident did not actually take place but added that this was an enactment of something that “could happen.” Suppose that one in 200 long-distance telephone calls is misdirected. What is the probability that at least one in five attempted telephone calls reaches the wrong number? (Assume independence of attempts.)

Ans:- Given that 1 Calls is misdirecting in 200 Calls

Probability of misdirecting Calls = 1/200

Probability of not misdirecting Calls = 199/200

No of calls =5

n=5 p=1/200 q=199/200

P(x)=Probability that at least one in five attempted telephone calls reaches the wrong number

Here, x=1

P(x)= P(x) = (nCx) (px) (qn-x) # nCr = n! / r! \* (n - r)!

P(1) = (5C1) (1/200)1 (199/200)5-1

P(1) = 0.0245037

1. Returns on a certain business venture, to the nearest $1,000, are known to follow the following probability distribution

|  |  |  |  |
| --- | --- | --- | --- |
| x | P(x) | E(X)=∑X \*P(X) | E(X2)=∑X2\*P(X) |
| -2,000 | 0.1 | -200 | 400000 |
| -1,000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  |  | 800 | 2800000 |

Ans :- inserted the new Table **E(X) = ∑ X \*P(X)**

**E(X2)=∑X2\*P(X)**

**Var(X)= E(X2) – {E(X)}2  & Std = √Var**

1. What is the most likely monetary outcome of the business venture?

$ 2000

1. Is the venture likely to be successful? Explain

Ans:- Successful Ventures are 1000,2000,3000

Probability = 0.2+0.3+0.1=0.6 as it is greater than 0.5 or 50% hence the venture is going to be successful

1. What is the long-term average earning of business ventures of this kind? Explain

Ans:- It is the average earning which is profit of $800

As Expected Value = E(X)=∑X \*P(X) = $800

1. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Ans:- Risk involved in the venture is the Var(X)

**Var(X)= E(X2) – {E(X)}2  = 2800000-8002**

**= 2160000 which is very High**

**Std = √Var = √2160000 ≈ 1470 Also very High**

**A**s Variance is very high the risk is also very high